

## Arquitetura de Computadores 2024/25

### TPC1

Este trabalho de casa consiste num exercício de programação **a ser realizado em grupo de, no máximo, dois alunos**. Pode esclarecer dúvidas gerais com colegas, mas a solução e a escrita do código devem ser estritamente realizadas pelos membros do grupo.

Os casos de plágio serão penalizados de acordo com os regulamentos em vigor. É possível usar ferramentas como o ChatGPT e o seu uso não necessita de ser reportado; *relembra-se que a nota obtida no trabalho é validada por perguntas específicas que aparecerão no 1º teste*.

**A entrega é feita via Mooshak até às 17:00 do dia 1 de Abril. Detalhes sobre a entrega serão divulgadas oportunamente, usando uma mensagem via CLIP.**

### Exercício a resolver

---

Um ficheiro é uma sequência de bytes armazenada no disco com um nome que é uma cadeia de caracteres. Relativamente ao seu conteúdo, um ficheiro pode ser classificado em duas categorias:

- **Texto:** contém apenas bytes correspondentes a algo que pode ser mostrado num terminal, a saber:
  - Bytes com os valores 9 (tab '\t'), 10 (carriage return, '\r') e 13 (line feed '\n')
  - Bytes entre 32 e 64: caracteres como espaço '\*', '!', e outros
  - Bytes entre 65 e 90: letras maiúsculas 'A'-'Z'
  - Bytes entre 91 e 96: mais caracteres, como por exemplo '\', '^', '[', ']'
  - Bytes entre 97 e 122: letras minúsculas 'a'-'z'
  - Bytes entre 123 e 126: mais caracteres como '{', '}', '|', '^', '~'
  - Bytes entre 160 e 255: caracteres acentuados e outros símbolos
- **Binário:** se contém algum byte que não faz parte da lista anterior.

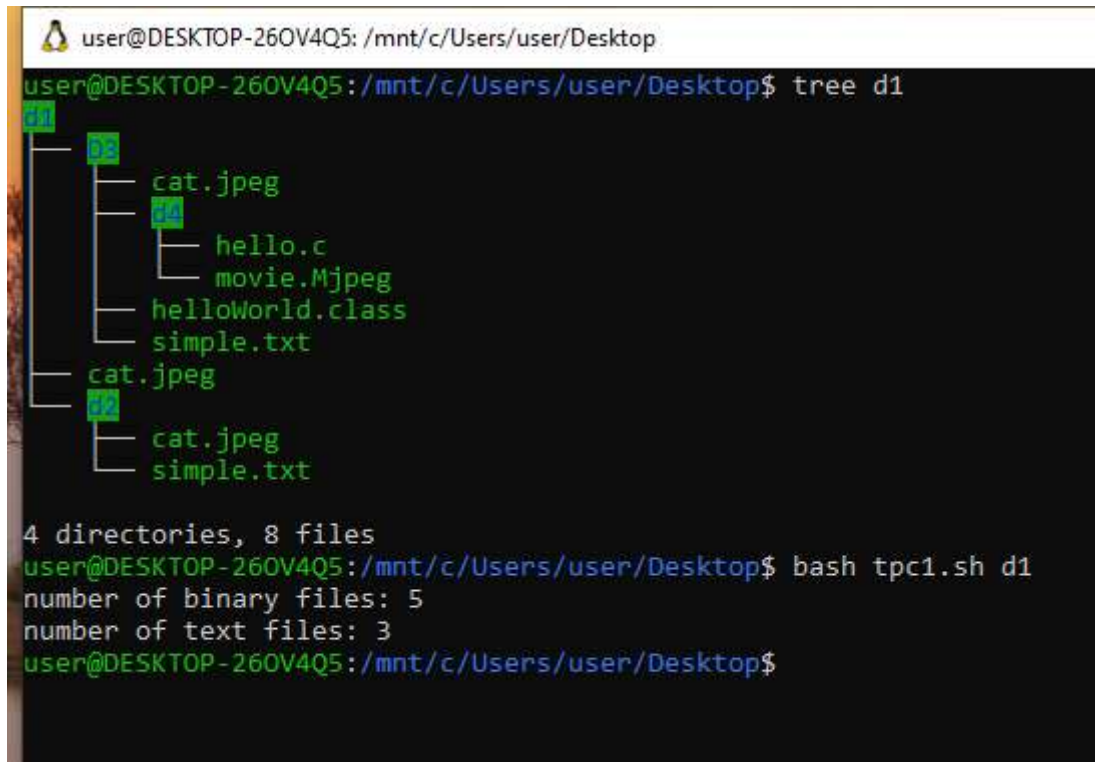
O trabalho prático consiste em criar dois programas:

- um programa em C chamado *isText* que aceita como argumento um ficheiro de entrada, e que retorna 1 se o ficheiro é um ficheiro de texto e 0 se é um ficheiro em binário, de acordo com as regras acima indicadas.
- um script em bash chamado *tpc1.sh* que encontra todos os ficheiros numa determinada sub-árvore do sistema de ficheiros e que invoca o programa *isText* para cada um dos ficheiros encontrados. De acordo com o valor retornado por *isText* vai atualizando dois contadores: um de ficheiros de texto e outro de ficheiros binários. No final, deve imprimir o número total de cada um dos tipos de ficheiros.

O ficheiro *tpc1.zip*, disponível no CLIP, contém um conjunto de pastas e ficheiros em que:

- *hello.c* e *simple.txt* são ficheiros de texto
- *cat.jpeg*, *helloWorld.class* e *movie.Mpeg* são ficheiros binários.

Na figura seguinte apresenta-se o conteúdo de uma sub-árvore do sistema de ficheiros, obtida através do comando *tree*<sup>1</sup>; segue-se os resultados da execução do script *tpc1.sh*.



```
user@DESKTOP-26OV4Q5: /mnt/c/Users/user/Desktop
user@DESKTOP-26OV4Q5:/mnt/c/Users/user/Desktop$ tree d1
d1
├── 02
│   ├── cat.jpeg
│   ├── 04
│   │   ├── hello.c
│   │   └── movie.Mjpeg
│   ├── helloWorld.class
│   ├── simple.txt
│   └── cat.jpeg
├── 02
│   ├── cat.jpeg
│   └── simple.txt
└── 02
```

4 directories, 8 files

```
user@DESKTOP-26OV4Q5:/mnt/c/Users/user/Desktop$ bash tpc1.sh d1
number of binary files: 5
number of text files: 3
user@DESKTOP-26OV4Q5:/mnt/c/Users/user/Desktop$
```

<sup>1</sup> Poderá ter de instalar o programa *tree*. No Ubuntu, terá de dar o comando *sudo apt-get install tree*.